Fundamentals of Fluctuation Spectroscopy IV:

Photon Counting Histogram Analysis

Don C. Lamb



Department of Physical Chemistry

Amplitude Fluctuations



Svedberg and Inouye, Zeitschr f. Physik Chemie 1911, 77:145-119

The measured probability function for detecting N photons in a time bin is a renormalization of the histogram of the photon counting data.

$$P_{\exp}(N) = \frac{freq(N)}{N_{Total \ photons}}$$

For a Poisson Distribution:

$$P(N) = \frac{\left\langle N \right\rangle^N e^{-\left\langle N \right\rangle}}{N!}$$

$$\left\langle \Delta N^2 \right\rangle = \left\langle N \right\rangle$$

 $\left\langle \Delta N^2 \right\rangle < \left\langle N \right\rangle$
 $\left\langle \Delta N^2 \right\rangle > \left\langle N \right\rangle$

super-Poissonian Statistics sub-Poissonian Statistics

Photon Counting Statistics

The number of detected photons from a constant intensity light source is governed by Poisson statistics

$$p(k,\langle k\rangle) = \frac{(\eta_E E)^k e^{-\eta_E E}}{k!} = \operatorname{Poi}(k,\langle k\rangle)$$

where: k is the number of detected photons

 $\langle k \rangle = \eta_E E$ is the average number of detected photons

 η_E is the detection efficiency

E is the energy impinging on the detector

e.g. A non-diffusing 500 nm fluorescent bead in the excitation volume:



Mandel's Formula

For a fluctuating intensity source, the photon counting distribution is given by Mandel's formula:

Mandel, Proc. Phys. Soc. (1958) 72:1037-1048

$$p(k,t,T) = \int_0^\infty \frac{(\eta_E E(t,T))^k e^{-\eta_E E}}{k!} P(E(t,T)) dE$$

where: P(E(t,T)) is the energy probability distribution

T is the integration time of the measurement

$$E(t,T) = \int_{t}^{t+T} I_D(t) dt$$

where: I_D is the intensity reaching the detector

Effect of binning:

For $T \rightarrow 0$: power fluctuations tract intensity fluctuations

For $T \rightarrow \infty$: intensity fluctuations average out,

$$\mathbf{p}(E) \rightarrow \delta(E - \langle E \rangle)$$

Choose *T* small enough to tract intensity fluctuations:

$$E(t) = I_D(t)T$$

$$p(k,t,T) = \int_0^\infty \frac{(\eta_I I_D(t))^k e^{-\eta_I I_D(t)}}{k!} P(I_D(t)) dI_D$$

where: $\eta_I = \eta_E T$

Diffusing Particle in a Confocal Volume

 I_D depends upon the position of the particle

The *PSF* gives the *measured fluorescence intensity* of a point particle at the position **r** within the probe volume

The intensity at the detector from a fluorophore at position **r** is given by:

$$I_D(\mathbf{r}) = \frac{I_{ex}^n \beta \ \overline{PSF}^n(\mathbf{r})}{n}$$

where n = number of absorbed photons per excitation β includes corresponding scale factors between excitation and detection intensity

We define the Molecular Brightness to be the measured intensity of a molecule at the center of the *PSF*:

$$\varepsilon = \frac{I_0^n \beta \eta_I}{n} = \frac{kQW^n(0)}{n}$$
$$\eta_I I_D(\mathbf{r}) = \varepsilon \overline{PSF}^n(\mathbf{r})$$
$$p^{(1)}(k;\varepsilon) = \int \frac{\varepsilon \overline{PSF}^n(\mathbf{r})}{k!} e^{-\varepsilon \overline{PSF}^n(\mathbf{r})} P(\mathbf{r}) d\mathbf{r}$$

$$p^{(1)}(k;\varepsilon) = \int \operatorname{Poi}(k,\varepsilon \,\overline{PSF}^n(\mathbf{r})) P(\mathbf{r}) d\mathbf{r}$$

PCH for Particles in a Box

The probability of detecting k photons from a single molecule in a box of volume V_0 is given by:

$$p^{(1)}(k;V_0,\varepsilon) = \int \operatorname{Poi}(k,\varepsilon \overline{PSF}^n(\mathbf{r})) P(\mathbf{r}) d\mathbf{r}$$
$$= \frac{1}{V_0} \int_{V_0} \operatorname{Poi}(k,\varepsilon \overline{PSF}^n(\mathbf{r})) d\mathbf{r}$$
The average count rate is:

The average count rate is:

$$\left\langle k \right\rangle = \frac{1}{V_0} \int_{V_0} \varepsilon \,\overline{PSF}^n(\mathbf{r}) \, d\mathbf{r}$$
$$= \frac{\varepsilon V_{PSF}}{V_0}$$



For multiple particles in a box:

$$p^{(2)}(k;V_0,\varepsilon) = \iint \operatorname{Poi}\left(k,\varepsilon \,\overline{PSF}^n(\mathbf{r}_1) + \varepsilon \,\overline{PSF}^n(\mathbf{r}_2)\right) P(\mathbf{r}_1) P(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2$$
$$p^{(N)}(k;V_0,\varepsilon) = \int_{1} \dots \int_{N} \operatorname{Poi}\left(k,\varepsilon \sum_{i}^{N} \overline{PSF}^n(\mathbf{r}_i)\right) \prod_{i}^{N} P(\mathbf{r}_i) d\mathbf{r}_i$$

The expression can also be written as a convolution:

$$p^{(2)}(k;V_0,\varepsilon) = \left(p^{(1)} \otimes p^{(1)}\right)(k,V_0,\varepsilon) = \sum_{r=0}^{\infty} p^{(1)}(r;V_0,\varepsilon) p^{(1)}(k-r;V_0,\varepsilon)$$
$$p^{(N)}(k;V_0,\varepsilon) = \left(p^{(1)} \otimes^{N \text{ times}} p^{(1)}\right)(k,V_0,\varepsilon)$$

PCH in an Open Volume

Particles can enter and leave the subvolume V_0

The probability of having N particles in the subvolume V_0 is given by:

$$p^{(1)}(N) = \operatorname{Poi}(N, \langle N \rangle)$$



The probability of observing *k* photons is given by the product of the probability of observing *k* photons with *N* particles in the volume multiplied by the probability of having *N* particles in the volume:

> $\Pi(k; \langle N_{PSF} \rangle, \varepsilon) = \sum_{N=0}^{\infty} p^{(N)}(k; V_0, \varepsilon) \operatorname{Poi}(N, \langle N_{PSF} \rangle)$ where $p^{(0)}(k; V_0, \varepsilon) = \delta(k)$

The average count rate is given by: $\langle k \rangle = \varepsilon \langle N_{PSF} \rangle$ Information available from analysis: $\varepsilon, \langle N_{PSF} \rangle$

Two key assumptions for PCH:



- 1) The molecule does not move significantly during a time bin
- The molecular brightness is constant in time and follows the spatial profile of the excitation volume (no reactions, photophysics, etc . . .)

PCH versus Concentration



Fits to Poisson distribution

At high concentration, Poission statistics dominate The super-Poisson nature of the distribution is seen in the tail of lower concentration measurements

PCH versus Concentration



PCH versus ε



Biophys J 77:553

The brighter the molecule, the more clearly the non-Poissonian statistics are observable

Bright, Slow Molecules



PCH with and without many dim molecules



Fluorescent Intensity Distribution Analysis

The differences between PCH and FIDA are:

- 1) Treatment of the excitation volume
- 2) Mathematical approach
- The number of photons for a single species in a *small* volume element is given by:

$$p_{dV_i}(k) = \sum_{m=0}^{\infty} Poi(m, \langle N \rangle) Poi(k, m \varepsilon \overline{PSF}(\mathbf{r}_i))$$

 $\langle N \rangle = c dV_i$

where:

$$p_{dV_i}(k) = \sum_{m=0}^{\infty} \frac{(cdV_i)^m}{m!} e^{cdV_i} \frac{\left(m\varepsilon \overline{PSF}(\mathbf{r})\right)^k e^{-m\varepsilon \overline{PSF}(\mathbf{r})}}{k!}$$

The total probability function is given by the convolution of all of the *small* volume elements

$$\Pi(k;\langle N\rangle,\varepsilon) = \sum_{m=0}^{\infty} p_{dV_i}(k-m;\langle N\rangle,\varepsilon) p_{dV_{j\neq i}}(m;\langle N\rangle,\varepsilon)$$
$$p_{dV_{j\neq i}}(m;\langle N\rangle,\varepsilon) = \sum_{n=0}^{\infty} p_{dV_{j\neq i}}(m-n;\langle N\rangle,\varepsilon) p_{dV_{k\neq i,j}}(n;\langle N\rangle,\varepsilon)$$

This leads to a large number of convolutions that is numerically 'clumsy and slow' to calculate

FIDA

The generating function is defined as:

$$G(v) = \sum_{n=0}^{\infty} p(n)v^n$$

The generating function has the following properties:

- 1) Under certain conditions, the generating function completely determines the distribution
- 2) The generating function of the sum of independent variables is the product of the generating functions (or sum of the logarithm of the generating functions)
- 3) Moments can be determined from the derivates of the generating function

For $v = e^{2\pi i\xi}$, $G(\xi)$ and p(n) are Fourier transform pairs

The generating function for a volume element dV_i is given by:

$$G(\xi; dV_i) = \sum_{n=0}^{\infty} p_{dV_i}(n) e^{2\pi i \xi n}$$

$$G(\xi; dV_i) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{(cdV_i)^m}{m!} e^{-cdV_i} \frac{(m\varepsilon \overline{PSF}(\mathbf{r}))^n e^{-m\varepsilon \overline{PSF}(\mathbf{r})}}{n!} e^{2\pi i \xi n}$$

$$G(\xi; dV_i) = \exp\left[cdV\left(e^{(e^{2\pi i \xi} - 1)\varepsilon \overline{PSF}(\mathbf{r})} - 1\right)\right]$$

Treatment of Volume in FIDA

The generating function is given by integrating over the dV:

$$G(\xi_i) = \exp\left[\int_V c dV \left(e^{\left(e^{2\pi i\xi} - 1\right)\varepsilon \overline{PSF}(\mathbf{r})} - 1\right)\right]$$

FIDA reduces the 3D integral over volume to a 1D integral.

For Example: 3D Gaussian.

Each concentric surface has the same brightness on the detector. Perform a transformation:





The 3D volume integral becomes a 1D integral

$$G(\xi_i) = \exp\left[\int_u^{\infty} c \frac{\pi w_r^2 w_z}{\sqrt{2}} \sqrt{u} du \left(e^{\left(e^{2\pi i \xi} - 1\right) e^{-u}\right)} - 1\right)\right]$$

FIDA defines the *PSF* empirically using:

$$dV = \left(a_1u + a_2u^2 + a_3u^3\right)du$$

where the coefficients a_1 , a_2 and a_3 are determined from the PCH/FIDA analysis of a known fluorescent standard

The photon counting distribution is determined from the discrete inverse Fourier transform of the generating function

$$p(n) = \sum_{\xi=0}^{\infty} G(\xi) e^{-2\pi i \xi n}$$

Distributions of Molecular Brightnesses

So far we have assumed each species has a well defined molecular brightness, ε_i

A distribution of molecular brightnesses can be fit to the PCH.





Warning! There are more parameters than data points. Criteria other than minimum χ^2 are needed to fit to the data.

Multiple Species

PCH distinguishes between different species via the molecular brightness, independent of the diffusion time.

$$p^{(N_1,N_2)}(k;V_0,\varepsilon_1,\varepsilon_2) = \int_{1} \dots \int_{N_1+N_2} \prod_{i=1}^{N_1+N_2} P(\mathbf{r}_i) d\mathbf{r}_i$$
$$\operatorname{Poi}\left(k,\varepsilon_1 \sum_{i=1}^{N_1} PSF(\mathbf{r}_i) + \varepsilon_2 \sum_{j=1}^{N_2} PSF(\mathbf{r}_j)\right)$$
$$\Pi(k;\langle N_1 \rangle,\varepsilon_1,\langle N_2 \rangle,\varepsilon_2) = \Pi(k;\langle N_1 \rangle,\varepsilon_1) \otimes \Pi(k;\langle N_2 \rangle,\varepsilon_2)$$



Background, Dark counts, scattered laser light, etc. can be treated as an additional species

Multiple Species

Mixture of 20 % rhodamine and 80 % coumarin



Measuring Labeling Efficiencies

PCH of alcohol dehydrogenase for (a) singly labeled protein and (b) a mixture of singly labeled and doubly labeled protein



PCH can also be used to investigate the amount of aggregation, formation of dimers, trimers, . . .

Resolvability







The small number of data points in the fit limitations the number of parameters one can reliably fit to.

2D PCH



With two channel detection, 2D PCH can be analyzed

2D PCH analysis provides additional data points for parameter determination with multiple species